

10/582420

1AP20 Rec'd PCT/PTO 09 JUN 2006

Docket No.: JEEKP103US  
Applicant: Zhao, et al.  
Title: GLYCINE-RICH PROTEINS, THEIR CODING GENES  
AND APPLICATIONS

I hereby certify that the attached patent application (along with any other paper referred to as being attached or enclosed) is being deposited with the United States Postal Service on this date June 9, 2006, in an envelope as "Express Mail Post Office to Addressee" Mailing Label Number EV763465319US addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Rebecca A. Bellas  
(Typed or Printed Name of Person Mailing Paper)

Rebecca A. Bellas  
(Signature of Person Mailing Paper)

10/582420

Glycine-Rich Proteins, Their Coding Genes and Application

SAP20 Rec'd PCT/PTO 09 JUN 2006

**Technical Field**

5 This invention relates to Glycine-rich protein (hereafter refer to Glyrichin), its coding gene and application, especially relates to Glyrichin of human and mouse origin and its coding gene, as well as its application in anti-bacterial area.

**Discussion of the Prior Art**

10 Over the past two decades, the medical circle has not discovered any new families of antibiotics, on the other hand, drug resistance has developed very rapidly. It is troubling that after the problem of *Staphylococcus aureus* resistance has just been overcome, American doctors have found a new bacterial strain that is resistant to the latest antibiotic, Vancomycin. The problem of drug resistant pathogenic bacteria is 15 increasingly threatening people's health. A completely novel type of antibiotic will be an effective way to resolve the resistant problem. As an active factor of organism's innate immunity, antimicrobial peptides are used by various organisms in defending against the invasion of pathogenic bacteria, which exist extensively in nature. The purified endogenous antimicrobial peptides have characteristics such as high antibacterial activity, broad antibacterial spectrum, many varieties, large scope for selection and low- 20 degree of resistant mutation of target bacterial strains. Therefore, anti-microbial peptides are considered to have wide prospects for application in the medical industry, food industry and agriculture.

25 Antimicrobial peptide refers to the polypeptide synthesized on ribosome through encoded genes. Their relative molecular masses are typically less than 10kDa. Due to their antibiotic activities, they are also called polypeptide antibiotics. Most antimicrobial peptides are stable under high temperature and their activities can be retained even after being heated for 10 to 15min at 100°C. Antibacterial peptides have strong tolerance to higher ion strength and extreme pH values. Most of the antibacterial peptides have 30 isoelectric points of >7 and show strong characteristics of positive ions. Meanwhile, some antimicrobial peptides have the capability to resist the hydrolysis by trypsases or pepsases. Additionally, studies have shown that antimicrobial peptides of different families have little homology in terms of sequence while different antibacterial peptides of the same family have high levels of homology in terms of sequence, which means 35 that their function is highly conserved. In addition to their antibacterial and antifungal activities, some antimicrobial peptides also have antiprotozoal, antiviral, and anti-tumor activities. Antibacterial peptides exist widely in living nature. Up to now, more than 700

kinds of polypeptide antibiotics have been found in bacteria, fungi, amphibians, pests, higher plants, mammals and even humans. Currently, preclinical feasibility studies of many antimicrobial peptides are being carried out. Additionally, studies on transgenic animals and plants with antimicrobial peptides, the application of antimicrobial peptides in food preservative, preservation of flowers, cosmetics, seed coating and animal feed additives are also ongoing.

Given the promising prospects for application in the medical, food and agricultural industries, the search for novel natural antimicrobial peptides in nature through various ways has become a worldwide priority.

10

### **Summary of the Invention**

The objective of this invention is to provide a Glyrichin, its coding gene and their antibacteria effects.

15 Glyrichin provided by this invention is selected (cloned) from at least one of the following glyrichin families:

- 1) Human Glyrichin (h Glyrichin) and mouse Glyrichin (m Glyrichin): protein having the amino acid sequence of the sequence 1 in the Sequence List or protein with antibacterial activities having sequence 1 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 1;
- 2) *Danio rerio* Glyrichin: protein having the amino acid sequence of the sequence 3 in the Sequence List or protein with antibacterial activities having sequence 3 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 3;
- 3) *Anopheles gambiae* Glyrichin: protein having the amino acid residue sequence of the sequence 4 in the Sequence List or protein with bacterial activities having sequence 4 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 4;
- 4) *Drosophila melanogaster* Glyrichin: protein having the amino acid residue sequence of the sequence 5 in the Sequence List or protein with bacterial activities having sequence 5 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20

amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 5;

5) *Caenorhabditis elegans* Glyrichin: protein having the amino acid residue sequence of the sequence 6 in the Sequence List or protein with bacterial activities having sequence 6 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 6;

10) *Caenorhabditis elegans* Glyrichin: protein having the amino acid residue sequence of the sequence 7 in the Sequence List or protein with bacterial activities having sequence 7 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 7;

15) *Schizosaccharomyces pombe* Glyrichin: protein having the amino acid residue sequence of the sequence 8 in the Sequence List or protein with bacterial activities having sequence 8 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 8;

20) *Saccharomyces cerevisiae* Glyrichin: protein having the amino acid residue sequence of the sequence 9 in the Sequence List or protein with bacterial activities having sequence 9 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 9;

25) *Arabiopsis thaliana* Glyrichin: protein having the amino acid residue sequence of the sequence 10 in the Sequence List or protein with bacterial activities having sequence 10 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 10;

30) *Plasmodium falciparum* 3D7 Glyrichin: protein having the amino acid residue sequence of the sequence 11 in the Sequence List or protein with bacterial activities having sequence 11 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to

20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 11;

5 11) *Plasmodium yoelii yoelii* Glyrichin: protein having the amino acid residue sequence of the sequence 12 in the Sequence List or protein with bacterial activities having sequence 12 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 12;

10 12) *Magnaporthe grisea* Glyrichin: protein having the amino acid residue sequence of the sequence 13 in the Sequence List or protein with bacterial activities having sequence 13 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 13;

15 13) *Neurospora crassa* Glyrichin: protein having the amino acid residue sequence of the sequence 14 in the Sequence List or protein with bacterial activities having sequence 14 in the Sequence List with 1 to 20 amino acid residues of it being deleted, inserted and / or substituted and with the 1 to 20 amino acid residues being added to the carboxyl terminal and / or amino terminal of sequence 14;

20 The preferred glyrichins hereof are human glyrichin and mouse glyichin.

The number of amino acid residues that undergo deletion, insertion and/or substitution as well as addition at the carboxyl terminal and /or amino terminal is preferred to be 1 to 10, more preferred to be 1 to 5 and most preferred to be 1 to 3.

25 The preferred substitutions hereof are shown in Table 1.

Table 1. List of substitutable amino acids of glyrichin family

Original residue	Representative Substitution	Preferred Substitution
Ala(A)	Val; Leu; Ile	Val
Arg(R)	Lys;Gln;Asn	Lys
Asn(N)	Gln;His;Lys;Arg	Gln
Asp(D)	Glu	Glu
Cys(C)	Ser	Ser
Gln(Q)	Asn	Asn
Glu(E)	Asp	Asp
Gly(G)	Pro; Ala	Ala

His(H)	Asn; Gln; Lys; Arg	Arg
Ile(I)	Leu; Val; Met; Ala; Phe	Leu
Leu(L)	Ile; Val; Met; Ala; Phe	Ile
Lys(K)	Arg; Gln; Asn	Arg
Met(M)	Leu; Phe; Ile	Leu
Phe(F)	Leu; Val; Ile; Ala; Tyr	Leu
Pro(P)	Ala	Ala
Ser(S)	Thr	Thr
Thr(T)	Ser	Ser
Trp(W)	Tyr; Phe	Tyr
Tyr(Y)	Trp; Phe; Thr; Ser	Phe
Val(V)	Ile; Leu; Met; Phe; Ala	Leu

Glyrichin described in this invention can be modified (the Primary Structure is usually not affected), such as acetylation, carboxylation, glycosylation and phosphorylation.

5 The coding gene of the Glyrichin hereof also is within the scope of this invention.  
 The coding gene of the human Glyrichin (hGlyrichin) hereof may contain DNA sequence of the SEQID No: 2 in Sequence List or DNA sequence having 90% homology with the DNA sequence in the SEQID No: 2 of the Sequence List, a DNA sequence corresponding to the amino acid residue sequence in the coding Sequence List 1, or nucleotide sequences hybridized with DNA sequence in the sequence 2 of Sequence List under high strict conditions.

The high strict conditions for the hybridization refer to a wash of the blotting membrane in  $0.1 \times$  SSPE (or  $0.1 \times$  SSC) and 0.1% SDS solution at  $65^\circ\text{C}$ .

10 The expression vectors, cell lines and engineered bacteria containing the coding gene of the Glyrichin hereof belong to the protection scope of this invention.

The Glyrichin in this invention hereof can be used for antibacterial purposes. Specifically, its application consists of the following aspects:

15 1) preparation of drugs for prophylaxis and/or treatment of human or livestock bacterial infectious diseases;  
 20 2) preparation of products for prevention and/or treatment of potentially infectious bacterial disease of various kinds of creatures;  
 3) production of transgenic organisms which can defend against diseases and pests;  
 4) preparation of derivatives, antagonists or ligands, or antibodies to Glyrichin.

### Brief Description of the Drawing

Fig. 1: The phylogenetic tree analysis of the Glyrichin family members in the comparison to the homology.

5 Fig. 2: Conserved structural domain of the Glyrichin family.

Fig. 3: Homology of the structural domain among different members of the Glyrichin family.

Fig. 4: Structural characteristics of the Glyrichin family.

10 Fig. 5A: Northern blot analysis results of the size and number of hGlyrichin transcripts using hGlyrichin gene as the probe.

Fig. 5B: The results of the in vitro transcription and translation of hGlyrichin.

Fig. 6A: The semi-logarithm growth curve of E. coli BL21 transformed with pET-22b (+) plasmid with or without the induction of 0.5mM IPTG.

15 Fig. 6B: The semi-logarithm growth curve of E. coli BL21 transformed with pET-22b(+)UBF and pET-22b(+)PTH with or without the induction of 0.5mM IPTG.

Fig. 6C: The semi-logarithm growth curve of E. coli BL21 transformed with pET-22b-hGlyrichin positive clone 1 and 8 with or without the induction of 0.5mM IPTG.

Fig. 6D: The semi-logarithm growth curve of E. coli BL21 transformed with pET-22b-hGlyrichin positive clone 1 with or without the induction of 0.5mM IPTG.

20 Fig. 7: PAGE electrophoresis result of the prokaryocytic expression product of hGlyrichin.

Fig. 8: The inhibitory effects of the expression products of yeast transformed by full-length or 5'-deleted hGlyrichin gene on the growth of E. Coli BL21.

25 Fig. 9: The growth curves of transformed BL-21 with pET-22b plasmids containing either fulllength, or 5'-deleted or 3'-deleted hGlyrichin gene with the induction of 0.5mM IPTG.

### Description of the Preferred Embodiment

The differentially expressed genes of mouse bone marrow stromal cells before and 30 after LTC (long term culture) are screened through the Suppression Subtractive Hybridization (SSH) method and 131 differentially expressed EST clones were harvested. It is verified through bioinformatics analysis that these clones represent 26 genes which are functionally known or partially known and 7 completely novel genes. Among them, 5 have a complete open-reading frame including the mGlyrichin gene 35 (GenBank number: AY028425) encoding amino acid sequence of mGlyrichin in Sequence 1.

After mGlyrichin was obtained, hGlyrichin with complete ORF sequence of 240bp is obtained through RT-PCR amplification using human embryonic liver mRNA as the template and primers designed based on the mGlyrichin sequence. The amino acid sequence of hGlyrichin is also obtained according to hGlyrichin nucleotide sequence.

5 The sequence is then specifically analyzed with means of bioinformatics and the homology comparison (multiple-sequence alignment) was performed among the family members.

The phylogenetic tree is analyzed within 12 sequences in Figure 1 in order to study the evolutionary relationships among Glyrichin family members. The evolution of 10 Glyrichin genes from lower to higher organisms and the genetic relationships among the family members are presented in Figure 1. The conserved Glyrich domains of Glyrichin family members (conserved amino acids are shown in Reverse) are presented in Figure 2. The sequence homology of Glyrich domains among different Glyrichin family members are presented in Figure 3, which demonstrates the close genetic relationships 15 of every two members. Results of electronic PCR showed that the hGlyrichin gene is located in the 20q11.21 area of human chromosome and consists of 3 exons encoding a small protein of 79 amino acids. Homology comparison with BlastP of amino acid sequences shows that this gene is very conserved in evolution, and exists in various species including fungi (e.g. *Neurospora crassa* and yeast), plants (e.g. *Arabidopsis thaliana*), *Plasmodium falciparum* 3D7, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Anopheles gambiae*, *Danio rerio*, mice and human, and forms the homologous 20 Glyrichin protein family. Homology comparison of the genes derived from various organisms shows that the highest homology, 100% homology, is demonstrated between hGlyrichin and mGlyrichin, 90% homology between hGlyrichin and *Danio rerio* 25 Glyrichin, 62% homology between hGlyrichin and *Drosophila melanogaster* Glyrichin, 46% homology between hGlyrichin and yeast Glyrichin. The genetic relationships between different family members are presented with a cladogram (Figure 1).

Among Glyrichin family members, the sequence characterizations are presented in Table 2.

30 Table 2. Sequence characterization of Glyrichin family members with sequences of amino acid residues in Sequence 1 and Sequence 3 to 14

Species Name	Number of			
	Length of Polypeptide (aa)	Electric		
		Isoelectric Point	Charges (pH7.0)	Glycine Content(%)
<i>Homo sapiens</i>	79	9.36	4.79	21.52

<i>Mus musculus</i>	79	9.36	4.79	21.52
<i>Danio rerio</i>	80	9.81	4.82	21.2
<i>Anopheles gambiae</i>	128	9.73	10.08	14.1
<i>Drosophila melanogaster</i>	79	9.66	4.82	20.3
<i>Caenorhabditis elegans</i>	145	10.79	12.01	14.48
<i>Schizosaccharomyces pombe</i>	107	10.04	4.05	13.08
<i>Saccharomyces cerevisiae</i>	112	9.74	5.04	13.27
<i>Plasmodium falciparum 3D7</i>	168	9.66	5.15	15.94
<i>Plasmodium Yoelii Yoelii</i>	167	9.16	4.15	15.28
<i>Magnaporthe grisea</i>	107	11.75	8.24	13.08
<i>Arabidopsis thaliana</i>	74	9.84	6.18	14.86

Human Glyrichin (hGlyrichin) is composed of 79 amino acid residues, with a molecular weight of 8.8kDa and isoelectric point of 9.36. The net positive electrical charge is 4.79 at pH 7 and the glycine content is 21.52%. The secondary structure and the characterizations of hGlyrichin have been analyzed by the inventor. Figure 4 is the 5 SignalP analysis result. The isoelectric points of the other members of Glyrichin family are all above 7 at pH 7 and they all have net positive electrical charge and strong hydrophobic domains. The above results show that the Glyrichin family has the typical structural features of known antimicrobial peptides.

The 13 members of Glyrichin family in this invention are from the species whose 10 genome sequencing has been finished, covering from fungi to human. Based on this, a person skilled in the art can simply infer that they might exist in the species whose genome sequencing has not yet been finished. The common feature in the structure of Glyrichin family members is that they all contain an area rich in glycine that is between 15 59 to 68 amino acids in length. This area almost covers the whole sequences of hGlyrichin and mGlyrichin. Besides the common area rich in glycine, Glyrichin existing in other species also contains other sequences (such as signal peptide etc.) at the amino terminal or carboxyl terminal of this conserved area, but they are not conserved (Figure 2 and 3). This shows that the conserved area rich in glycine might represent basic 20 biological activities of Glyrichin family members and that the redundant sequences were gradually eliminated in the process of biological evolution, making the size of molecule smaller and the structure more compact.

The coding genes of Glyrichin in this invention can be used to express or produce recombinant Glyrichin protein through routine recombinant DNA technology. The method of expression or production is as follows:

25 (1) Use the recombinant expression vectors containing the Glyrichin coding gene

mentioned above to transform host cells;

(2) Cultivation of the host cell;

(3) Separate and purify protein from the medium or cells.

The starting vectors used to construct the recombinant expression vector can be the

5 known vectors in this area, such as bacteria plasmid, bacteriophage, yeast plasmid, plant cell virus, mammal cellular virus including adenovirus or other vectors. In general, those plasmids and vectors that can replicate and express stably in host cells are applicable.

10 An important feature of expression vectors is that they usually contain a replication origin, promoter, marker gene and translational control component. The methods known to those skilled in the art can be used to construct expression vectors containing Glyrichin DNA coding sequence and appropriate transcription/ translation control signals.

15 Those skilled in the art know how to choose appropriate vectors, promoter, enhancer and type of host cells. The transformation of host cells with the above mentioned recombinant expression vectors may be conducted with the routine techniques known to those skilled in the art. The obtained transformants can be cultured with routine methods and Glyrichin protein can be expressed intracellular or on cell membrane, or can be excreted to the outside of the cells. If necessary, the recombinant 20 protein can be separated and purified with various separation methods according to its physical, chemical and other characteristics. Other drugs, such as antibiotics including penicillin can be used in combination with Glyrichin protein in this invention.

25 In this invention, terms including “Glyrichin”, “Glyrichin family member”, “Glyrichin protein ” or “Glyrichin polypeptide” are interchangeable, and all of them refer to protein or polypeptide with amino acid sequence (sequence 1 or sequence 3-14 in the Sequence List 1) of the family members of natural antimicrobial peptide Glyrichin. They include those family members of natural antimicrobial peptide Glyrichin with or without starting methionine as well as Glyrichin protein with or without signal peptide.

30 The term “Glyrichin structural domain” refers especially to a segment of the most conserved polypeptide in Sequence 1 to 13 in the Sequence List with 59 to 68 amino acid residues. Their isoelectric points are greater than 7.0 and they have net positive electrical charges at pH 7.0. The Glycine content in this conserved domain is the highest of the 20 amino acids and there is at least one hydrophobic region. Their sequences and 35 structural features are presented in Figures 3 and 4 and Table 2.

The term “Glyrichin family” refers specifically to polypeptides that have a Glyrich

domain and the homology of amino acids in this domain is 30% or higher, including the 13 polypeptides mentioned in this invention and other homologous polypeptides not reported. From the view of evolutionary biology, they are from the same origin.

5       Fusion proteins containing Glyrichin polypeptide or its segments are also involved in this invention. Besides polypeptides of almost full length, soluble segments of Glyrichin polypeptide are also included in this invention. Usually, the segments contain a certain number of consecutive amino acid sequences in the Glyrichin polypeptide sequence. It should be understood that the polypeptides in this invention are not limited to the representative polypeptides that are cited above.

10       The term “modification” (usually the primary structure is not changed) includes: in vivo or in vitro chemical modification of polypeptides such as acetylation or carboxylation. Modification also includes glycosylation, for example, the polypeptides formed in the process of synthesis and processing or further processing through glycosylation. The modification can also be accomplished through exposure of 15 polypeptide to enzymes that catalyze glycosylation (e.g. glycosylase or deglycosylase of mammals). Modified forms also include the sequence containing phosphorylated amino acid residues (such as phosphoserine and phosphothreonine) as well as polypeptide whose proteolytic resistance or solubility is improved through modifications.

20       The function of the Glyrichin gene family has been representatively studied by the inventor using hGlyrichin and mGlyrichin as examples. Results show that hGlyrichin and mGlyrichin have antibacterial activities, which are extensively expressed in various tissues of humans and mice. The antibacterial peptides exist naturally and are likely to be helpful in the medical industry and all fields that require the prevention of bacterial infection. This is also the basic function and activity of Glyrichin family members.

25       This invention is further explained below using the cloning, expression, purification and antibacterial activity tests of hGlyrichin as an example. It should be understood that these examples should only be used to explain this invention instead of confine it. The experimental methods without stated conditions usually use standard conditions such as the conditions described in Molecular Cloning: Laboratory Brochure 30 (New York: Cold Spring Harbor Laboratory Press 1989) by Sambrook *et al.* or the conditions recommended by manufacturers.

#### Example 1: Acquisition and analysis of hGlyrichin gene

##### 1. Acquisition of the complete ORF sequence of hGlyrichin gene

35       The following pair of primers is designed based on the mGlyrichin gene and RT-PCR reaction is conducted using human fetal liver RNA as the template.

5'primer Pa: 5'-CGATGCCGGTGGCCGTGGGTCCCT-3'

3'primer Pb: 5'-TTAGCATCGTATGCCCATTCGA-3'

PCR reaction system: PCR reaction system is done according to standard molecular biology techniques. The amplification conditions of PCR are as follows: 4 min at 94°C for 1 cycle; 40 sec at 94°C, 50 sec at 60°C, 1 min at 72°C for 30cycles; 7 min at 72°C for 1 cycle. PCR products are purified using a Wizard PCR Purification kit (purchased from Promega), PCR product is sub-cloned into pGEM-T vector with T4 ligase to form pGEM-T/hGlyrichin, and then transformed into JM109 for sequencing. Results show that the Glyrichin gene has the nucleotide sequence of sequence 2 in the Sequence List and the coding sequence of amino acid residues in the coding Sequence List 1.

## 2. Northern blot analysis

Four cell lines from human tumors, HepG2, HeLa, Jurket and HEK293, are cultured and the total RNAs are extracted, respectively, with a Wizard Plus mRNA purification kit (purchased from Promega). Twenty mg of each total RNA is sampled and separated on a 1.2% formaldehyde denaturing agarose gel, and then transferred onto a Hybond-N+ nylon membrane. It is then labeled with a Prime-a-Gene kit from Promega using the complete ORF of hGlyrichin as the probe. The hybridization results are presented in Figure 5A, which shows that the hGlyrichin gene is expressed in the four tumor cell lines from different tissue sources, indicating that hGlyrichin may be an extensively expressed natural antibacterial peptide. The results also demonstrate that there is only one transcript with a size of approximately 600bp.

## 3. In vitro transcription and translation test

The full length ORF of hGlyrichin gene was obtained though PCR using pGEM-T/hGlyrichin plasmid as the template and the following primers:

5'primer: 5'-CGGGATCCCGATGCCGGTGGCCGTGGGTCCCT-3'

3'primer: 5'-GCTCGAGTTAGCATGGATGCCCATCC-3'

The regular (refer to short protocols in molecular biology) PCR reaction system is adopted and the conditions of PCR amplification are as follows: 4 min at 94°C for 1 cycle; 40 sec at 94°C, 50 sec at 60°C, 1 min at 72°C for 30cycles; 7 min at 72°C for 1 cycle.

The amplification products of PCR are then incised with BamHI and SacI; pT7 vectors (purchased from Promega) are incised with the same enzymes and then linked with T4 DNA ligase to construct pT7-hGlyrichin plasmid. The constructed plasmids are then transcribed and translated in the following system:

25  $\mu$ l system:

Rabbit Reticulocyte lysate (RRL): 12.5  $\mu$ l  
Reaction buffer: 1.0  $\mu$ l  
Amino acids mixture: 0.5  $\mu$ l  
<sup>35</sup>S-Met: 1.0  $\mu$ l(50  $\mu$ l)  
5 DNA 2.0  $\mu$ l(0.5  $\mu$ g/ $\mu$ l)  
RNase inhibitor: 0.5  $\mu$ l  
T7 DNA polymerase: 0.5  $\mu$ l  
Deionized water: 7.0  $\mu$ l  
Total volume: 25.0  $\mu$ l

10

The system is incubated at 30°C for 90 minutes. Five  $\mu$ l of sample is taken after the reaction is finished and then 10 $\mu$ l loading buffer is added into the sample to perform the SDS-PAGE. The gel is fixed in a fixing solution for 30 minutes, immersed in the desiccant for 5 minutes, and dried over night on a drying rack. Autoradiography is then 15 performed and the film is exposed at -20°C for 24 hours. The results are presented in Figure 5B. The results demonstrate that the size of the protein encoded by the hGlyrichin gene is about 8.8 kDa after *in vitro* translation, which is the same as the results deduced by theory. This indicates that the gene can be transcribed and translated normally *in vitro*. The reagents used in the reaction system are from the TNT kit from 20 Promega, and <sup>35</sup>S-Met is purchased from Amersham Bioscience Corp. In Figure 5B, Lane 1 is the molecular weight reference, Lane 2 is the UBF of unrelated genes (GenBankUBF-f1 AF294842; Chinese Journal of Applied Physiology, 20:66, 2004), Lane 3 is the translation products of hGlyrichin gene and Lane 4 is the translation 25 products of the positive control provided in the kit.

25

Example 2: The inhibitory effects of hGlyrichin (mGlyrichin) on the growth of *Escherichia coli* BL21 transformed with different plasmid containing hGlyrichin gene or other genes with or without IPTG induction.

1. The construction of pET-22b-hGlyrichin

30 Use pGEM-T/hGlyrichin as the template, and 5'-  
GGAATTCCATATGCCGGTGGCCGTG GGTC-3' as well as 5'  
CCGCTCGAGTTAGCATCGGATCCCATC-3' as primers, amplify hGlyrichin gene  
through standard PCR amplification. The PCR reaction system (excluding primer) and  
reaction condition are the same as those of step 3 in Example 1.

35 Use NdeI and XhoI enzymes to cleave, then run a gel, purify the amplification products of hGlyrichin, and sub-clone into expression vector pET-22b (+) that has been

cleaved by NdeI and XhoI enzymes in the presence of T4 DNA ligase. Transform *E. coli* BL21 and then identify the positive clones 1 and 8 (containing plasmid pET-22b-hGlyrichin) by means of enzyme cleavage.

2. Construction of Control Plasmids pET-22b(+)-UBF and pET-22b(+)-PTH

5 Introduce the UBF gene (GenBankUBF-fl AF294842; Chinese Journal of Applied Physiology, 20:66, 2004) and the PTH gene (the 34 peptide gene sequence of parathyroid hormone) (GenBankNM000315) into the multiple cloning sites of pET-22b(+) vector respectively, thus obtain the control plasmids pET-22b(+)-UBF and pET-22b(+)-PTH.

10 3. Tests of growth inhibition

Select the positive clones and inoculate them into Liquid LB Culture Medium (containing 50 $\mu$ g/ml ampicillin). Incubate them for 12 hours in the shaker, 250rpm at 37°C. Reinoculate them into a test tube in 1:100 (v/v) ratio and continue to incubate them until the OD value is 0.03. Then add IPTG into each tube to a final concentration is 0.5mM and add the same volume of PBS to the negative control tube. Culture the positive and negative groups in a shaker, 250rpm, at 30°C; then take out 1ml of culture from each tube every 45 minutes and take an OD600 measurement in succession for ten times or more. Draw a Growth Curve with time as the horizontal axis and the logarithm of OD600 as the longitudinal axis. The results are shown in Figures 6A, 6B, 6C and 6D.

15 20 Figure 6A showed that the bacterial growth curves basically coincide with each other whether IPTG is added to the bacterial strain transformed with pET-22b(+) blank Vector or IPTG is not added. Figure 6B shows that there is no influence on the growth of bacteria transformed with plasmids containing UBF or PTH gene regardless of whether IPTG is added or not. Figure 6C and 6D shows that the growth curves of the bacteria transformed with plasmid containing hGlyrichin are obviously different between IPTG induced or uninduced. The bacterial growth is obviously inhibited when the bacteria transformed with hGlyrichin-containing gene is induced by IPTG. From Figures 6C and 6D, it appears that as the time of culture prolongs, the IPTG is depleted and the bacterial growth resumes again. However, there is no such phenomena when the bacteria are transformed with unrelated gene-containing plasmids. These results fully show that the growth inhibitory effects are exerted by Glyrichin protein itself.

30 Example 3: The purification of expression products of hGlyrichin and its inhibitory effects on the growth of bacteria

35 Primers designed are upstream 5'-  
CGGGATCCGATGCCGGTGGCCGTGGTCCCT-3' and down stream 5'-

5 GGAATTCTTAGCATCGTATGCCATTCCA-3' in accordance with the cDNA sequence of mGlyrichin. PCR amplification is performed using human fetal liver mRNA after reverse transcription as the template. The PCR reaction system (except primer) and reaction conditions are the same as those of Step 1 in Example 1. After 10 purification, the purified PCR products are sequenced and digested with restriction endonuclease BamHI and EcoRI, and then inserted into prokaryotic expression vector pGEX4-4T2 (construct from Pharmcia) that have been digested by the same restriction endonucleases in the presence of T4 DNA ligase. Transform *E. coli* JM109, induce expression *in vitro* for 5 hours with 0.5mM IPTG, and collect the bacteria. Re-suspend 15 the bacteria with PBS and carry-out several freeze-thaw cycles repeatedly. Centrifuge at 4°C, 12000rpm for 20 minutes and then take some of the supernatant for SDS-PAGE. The results are shown in Figure 7, which show that the fusion protein of GST-Glyrichin is expressed in the supernatant and its molecular weight is 34kD. Lane 1 Figure 7 is protein molecular weight standards, Lane 2 is the induced expression products of GST- 20 hGlyrichin fusion protein, Lane 3 is the non-induced expression products of GST- hGlyrichin fusion protein, and Lane 4 is the induced expression products of GST protein, and Lane 5 is the non-induced expression products of GST protein.

25 Purify the GST-Glyrichin fusion protein obtained by induced expression with a Sepharose 4B GST purification column (Pharmcia Inc.), cleave with Enterokinase (Roche) to obtain free hGlyrichin protein. The antibacterial activities (against *E. coli* DH5 $\alpha$  and *Bacillus subtilis* DB430) of GST-Glyrichin fusion protein and free hGlyrichin protein are assayed by means of a 96-well plate and minimal inhibition concentration (MIC) is determined. When the antibacterial activity is assayed, dilute the bacterial liquid to a concentration between  $10^4$ - $10^5$ CFU/ml, inoculate the bacterial liquid into the 96-well plate, 80  $\mu$ l per well. Dilute the polypeptides at a certain ratio 30 and add 5  $\mu$ l per well. Cultivate the 96-well plate at 37°C for 12 hours, then measure the OD600 value with a UV/Vis spectrophotometer. The results show that the GST-Glyrichin fusion protein has no inhibitory effect on the growth of bacteria while free Glyrichin has inhibitory activity on the growth of bacteria at relatively low concentration. Free Glyrichin protein has antibacterial activity and can inhibit the growth of Gram+ bacteria and Gram- bacteria at relatively low concentrations (Table 3).

Table 3. MIC of free hGlyrichin protein

	Name of bacterial strain	MIC Concentration (ug/ml)
G- Bacteria	<i>E. coli</i> DH5 $\alpha$	0.2
G+ Bacteria	<i>Bacillus subtilis</i> DB430	3

Example 4: The yeast expression of hGlyrichin and the assay of the antibacterial activity of the expression products

PCR primers are as follows: the full-length upstream primer: 5'-AGGAATTTCATG  
5 CCGGTGGCCGTGGGTCCCTAC-3' ; 5'-deleted upstream primer: 5'-  
AGGAATTCACTGGGCTT CGTGATGGGTTGC-3; full-length downstream primer: 5'-  
AAGGAAAAAAAGCGGCCGCTTA GCATCGGATGC CCATCCCAATG-3'). PCR  
amplification is carried out respectively using pET-22b plasmid containing the full  
length of hGlyrichin as the template, in the presence of either a pair of upstream and  
10 downstream primers for full-length hGlyrichin or 5'-deleted upstream primer and full  
length downstream primer for 5'-deleted hGlyrichin. The 5'-deleted gene refers to a  
deletion of 60 bases at the 5'-terminal of hGlyrichin. The PCR reaction system (except  
primer) and reaction conditions are the same as those in step 3 of Example 1.

Insert the full-length and 5'-deleted fragment of the hGlyrichin gene obtained  
15 through PCR amplification, respectively, into pPIC9K Yeast Expression Plasmid  
(Invitrogen) between the EcoRI and NotI sites in the presence of T4 DNA ligase and  
transform *E. Coli* BL-21 engineering bacteria. Screen and obtain the constructed  
plasmids containing either full-length or 5'-deleted hGlyrichin. Linearize the highly  
purified plasmids with Sall endonuclease, perform electro-transforming method to  
20 transform GS115 yeast, and obtain positive clones by screening with a MD plate  
containing G418 (50 $\mu$ g/ml). Inoculate the positive clone into the culture media  
containing 5 ml BMGY and cultivate it in a shaker at 30°C until the OD600 is 2.0 to 6.0.  
Then dilute with BMMY culture media (containing methanol to a final concentration  
1%) until the OD600 is 1.0 and continue culturing at 30°C. For the induction, add  
25 methanol every 24 hours until the final concentration is 0.5%, pipette 1ml of culture at  
different time intervals post induction into Eppendorff centrifuge tubes, centrifuge at  
16000 rpm, and collect the supernatant for activity assay using the Kirby-Bauer disk  
diffusion test. The test is performed by applying the bacteria (*E. coli* BL-21) on to a LB  
plate containing 1% agar and then placing a filter paper of 2 mm diameter at the center  
30 of each small block. Drop 15 $\mu$ l of different cloning supernatants onto the filter paper,  
then drop the same volume after 20 minutes; drop three times in total. Culture the LB  
plate at 37°C for 4 more hours until the antibacterial rings can be observed, accordingly,  
the positive clones are determined. In this test, use 10 $\mu$ l ampicillin at the concentration  
35 of 100mg/ml as the positive control. After screening, positive clones expressing either  
full-length hGlyrichin (L4 clone) or 5'-deleted hGlyrichin (clone of S2, S12 and S5) are  
obtained. The inhibitory effects of L4, S2, S12 and S5 clones are shown in Figure 8,

which shows that the supernatant of positive clones (L4, S2, S12 and S5) expressing either the full-length or 5'-deleted hGlyrichin gene markedly inhibit the growth of *E. coli* BL-21 on the surface of the agar plate. The experimental results described above verify that the yeast expressing system can express hGlyrichin protein. Furthermore, its expression products, full-length or 5'-deleted, have antibacterial activities. These results lay a very solid basis for the future discovery of the definite structure responsible for the activity of this gene. According to the experimental results described above, persons skilled in the art can easily screen for the smallest hGlyrichin protein molecule with similar activity through similar methods.

10

#### Example 5: Studies of the structure-activity relationships of hGlyrichin gene

Growth comparison of different transformants transformed with plasmids containing various deleted hGlyrichin gene

The inventor designed a series of primers according to the sequence of hGlyrichin:

15

Primer 1: 5'-GGAATTCCATATGCCGGTGGCCGTGGTC-3'

Primer 2: 5'-GGAATTCCATATGGGCTTCGTGATGGGTTGC-3'

Primer 3: 5'-CCGGCTCGAGTTAGAATGTGCCAAAGGT-3'

Primer 4: 5'-CCGCTCGAGTTAGCATCGGATGCCATC-3'

20

In order to obtain different length fragments of the hGlyrichin gene, different combinations of primers are used to create full length (primer 1 and 4 combination), 5'-deleted sequence (primer 2 and primer 4 combination), and 3'-deleted sequence (primer 1 and primer 3 combination) by PCR amplifications using pET-22b-hGlyrichin as the template. The PCR reaction systems (primer excluded) and reaction conditions are the same as those in Step 3 of Example 1. A series of plasmids containing hGlyrichin gene fragments of different lengths are constructed by inserting each fragment deriving from PCR into pET-22b vector by using the method in Example 2. After transformation of *E. Coli* BL-21, screening of clones and identification, the transformed bacterial clones containing different constructed plasmids are used to induce the expression of different fragments of hGlyrichin using the method in Example 2 in order to observe the effect of the full length and different deleted fragments on the growth of transformed bacteria. The results are shown in Figure 9, which show that there are no inhibitory effects when *E. coli* BL-21 is transformed with vacant vector and expression induced with or without IPTG. When transformed with plasmid containing full-length or deleted fragment of hGlyrichin, the growth of bacteria is obviously inhibited within 5 hours after IPTG induction; however, the growth is not affected if IPTG is not added. In the Figure 9, the x-axis is the time (hr) (in the bracket) after IPTG induction and the y-axis is the value of

OD600 absorbance. The above results demonstrate that some nucleotide sequences at the 5'-terminal and 3'-terminal of the hGlyrichin gene are not necessary for its antibacterial effect. The (-) in Figure 9 indicates that IPTG is not added, while (+) indicates that 0.5mM IPTG is added.

5

#### **Industrial Applicability**

From the experimental results, the inventor for the first time discovers a Glyrichin family containing conserved sequences with antibacterial activity. Glyrichin can be applied in medical fields and all those fields where antibiotics are needed. Thus it has

10 very broad application prospects.